感觉记忆中含语义对象的视觉表征*

胡晶晶[1] 徐昊骙[1] 曹立人[1]

（[1]浙江大学心理与科学行为系，杭州 310028）

摘要 感觉记忆能够以特征形式构建对视觉对象表面信息的表征，然而对于具有重要意义的潜在抽象信息，尚未有研究考察其是否影响感觉记忆的视觉表征。结合掩蔽技术与部分报告法范式考察了感觉记忆的视觉表征中是否包含语义信息。以阿拉伯数字作为目标，实验 1 与实验 2 分别采用正立与倒置的汉字大写数字作为掩蔽刺激，比较其与白噪音刺激的掩蔽效应。结果显示，汉字大写数字的掩蔽效应均强于白噪音刺激，且正立汉字大写数字的掩蔽效果较倒置汉字大写数字更强。实验 3 采用汉字大写数字与汉字常用字分别作为掩蔽刺激，结果显示汉字大写数字的掩蔽效果强于汉字常用字。上述结果显示，刺激间的语义相似性带来了视觉表征冲突，表明感觉记忆的视觉表征中包含了刺激的语义信息。

关键字：感觉记忆；视觉表征；语义信息；视觉掩蔽

1 前言

日常生活中充斥着丰富的信息，视觉系统对信息的有效加工是人类与环境实现交互的重要基础。由于人类的视觉加工资源极其有限(Luck & Vogel, 1997; Raymond, Shapiro, & Arnell, 1992; Rensink, 2002)，视觉通常需要对信息进行选择性加工(Wahn & König, 2017)。然而，对信息的选择往往难以在信息呈现阶段完成，因此视觉系统需要将大量信息短暂地存储于感觉记忆(Coltheart, Laming, Routh, & Broadbent, 1983; Smithson & Mollon, 2006)。Sperling(1960)采用部分报告法首先揭示了感觉记忆的存在。当被试根据随机呈现的后线索报告部分目标时，其存储容量远高于视觉短时记忆。进一步研究发现，感觉记忆的维持时间较短，其存储的信息随时间迅速消退(Averbach & Coriell, 1961; Sperling, 1960)。通过视觉系统的选择机制，感觉记忆中的少量信息得以巩固，并存储于短时记忆中(Gegenfurtner & Sperling, 1993; Öğmen & Herzog, 2016)。

视觉表征是视觉研究的核心问题之一(Marr, 1982)，数十年来，感觉记忆的视觉表征得到大量研究的关注。早期研究认为，视觉感觉记忆以简单快照（icon）的形式构建对

收稿时间：2018-9-19

* 通讯作者：曹立人，E-mail: caoliren2000@163.com

物理世界的表征(Dick, 1974; Neisser, 1967; Sperling, 1963)。采用掩蔽技术的研究发现，相继呈现的多个视觉刺激，会产生相互融合或覆盖(Enns & Di Lollo, 2000; Scheerer, 1973)，时间间隔较短时产生融合(Schultz & Eriksen, 1977)；间隔较长时产生覆盖(Turvey, 1973)。近年来一些研究对快照式的表征观点提出了补充。Bhardwaj, Mollon 和 Smithson(2012)的研究显示，当目标刺激（如：阿拉伯数字 2856）与掩蔽刺激（如：阿拉伯数字 3491）外形结构相似时，被试对目标刺激的记忆绩效大幅下降。Pinto，Sligte，Shapiro 和 Lamme(2013)的研究也发现，与不具有朝向信息的掩蔽刺激（如：圆形）比较，带有朝向信息的掩蔽刺激（如：带朝向的长方形）显著降低了对目标刺激（如：不同朝向的长方形）朝向信息的记忆绩效。研究者们认为，被试对目标刺激的记忆绩效下降，是由目标与掩蔽刺激的视觉表征产生了冲突所致，由此可推论表征中含有刺激对象的特征信息。上述研究表明，快照或许并非感觉记忆唯一的视觉表征形式，视觉感觉记忆能够以更为灵活的特征形式表征刺激对象不同类型的信息。

如前文所述，现有研究已证明感觉记忆能够以特征形式构建刺激对象的视觉表征，然而此类研究主要针对的是刺激对象的表面信息，尚未见相关研究直接考察刺激对象背后的潜在信息对感觉记忆中视觉表征的影响。视觉的本质是对客观世界的主观理解(Rescorla, 2015)，以自上而下和自下而上相结合的方式加工刺激对象(Wolfe, Butcher, Lee, & Hyle, 2003)。有研究表明，成人能够从简单的视觉刺激中识别语义概念(沈模卫，潘善会，陈新，1998)、生物运动(Neri, Luu, & Levi, 2006)、社会关系(Scholl & Gao, 2013)等抽象信息；即使是婴幼儿，也同样能够理解视觉对象背后的概念和意图信息(Schlottmann & Ray, 2010; Yin & Csibra, 2015)。认知建模的相关研究指出，人类的知识经验以先验的方式参与到视觉加工中，与刺激呈现的表面信息共同完成对刺激背后潜在信息的逆向推理(Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017)。来自视觉各领域的研究均指出，视觉系统不仅描述刺激对象所呈现的表面信息，还能够识别和理解刺激对象潜在的抽象信息。考察刺激对象中抽象信息对感觉记忆视觉表征的影响，能够更为完整地揭示视觉感觉记忆中信息的表征特点，同时有助于进一步理解感觉记忆与人类知识经验在认知加工中的交互过程。以往涉及早期视觉阶段对抽象信息加工的研究，往往采用字词及其语义概念信息作为研究对象。字词作为语言的符号化形式，其外在的表面信息与潜在的语义信息具有紧密的绑定关系，有利于确保研究的有效性。大量研究证明语义信息能够在视觉加工早期被提取(Perfetti & Bell, 1991; Perfetti & Zhang, 1991; 沈模卫，高涛，丁海杰，2004)，那么这些语义信息能否在感觉记忆的视觉表征中得以体现，并表现为视

觉表征的特性呢？本研究将直接针对该问题加以探讨。

研究采用部分报告法测量视觉感觉记忆的存储绩效，同时结合视觉掩蔽技术，构建语义信息不同的掩蔽刺激，考察掩蔽刺激类型对记忆绩效的影响。具体来说，借助汉字大写数字与阿拉伯数字在语义上相似、形态上不同的特点，构建具有语义相似性的掩蔽刺激，并测量其掩蔽效应与不含语义信息/含不同语义信息的掩蔽刺激之间的差异。若感觉记忆的视觉表征中包含了对象的语义信息，则带有相似语义信息的掩蔽刺激，其表征可能干扰对目标刺激的表征，从而导致语义相似的掩蔽条件中目标刺激记忆绩效下降。本研究包含三个实验，实验 1 以阿拉伯数字作为目标刺激，比较正立的汉字大写数字和随机噪点两类掩蔽条件下的记忆绩效。由于汉字大写数字在语义上与阿拉伯数字相似，若该条件下的记忆绩效较随机噪点条件更低，则表明视觉感觉记忆中形成了包含语义信息的表征。考虑到汉字大写数字与随机噪点的形态差异较大，实验 2 采用倒置的汉字大写数字作为掩蔽刺激，使其物理特征与正立汉字大写数字基本相同，但破坏其语义信息 (Kao, Chen, & Chen, 2010)。若倒置操作改变了实验 1 中所观察到的掩蔽效应差异，则进一步证明不同掩蔽条件下记忆绩效的差异确实是由掩蔽刺激与目标刺激的语义相似性造成的。与白噪音和倒置的汉字大写数字相比，正立的汉字大写数字具有更强的语义性，这可能导致实验 1 与实验 2 的差异性源于掩蔽的语义性强弱，而非语义的相似性。为进一步考察上述问题，实验 3 采用汉字常用字和正立的汉字数字作为掩蔽刺激，两者具有相同强度的语义信息，但与目标刺激的语义相似性不同。如果实验 3 中，正立汉字大写数字的掩蔽效果较汉字常用字更强，则表明掩蔽刺激与目标刺激的语义相似性会造成掩蔽效应的差异，即视觉感觉记忆中对刺激的表征包含语义信息。

2 实验 1

2.1 方法

2.1.1 被试

共 24 名被试参与了实验 1（男性 10 名，女性 14 名），平均年龄为 21.29±2.51 岁。所有被试均为在校大学生，母语均为汉语，均视力(或矫正视力)正常。被试在实验前不清楚实验目的。

2.1.2 刺激材料

实验所用刺激通过 19 英寸的 CRT 显示器呈现（型号：DELL，P992），显示器的刷

新率被设定为 85Hz，分辨率为 1280×962 像素。实验程序采用 Matlab 的工具包 PsychotoolBox 编写(Brainard, 1997; Pelli, 1997)。实验在密闭、昏暗的房间中进行，主要的光源来自显示屏。被试通过计算机键盘来记录反应。

目标刺激为以 3×4 方阵排列在屏幕中央的随机阿拉伯数列，字体为 Arial，在 60cm 的视距下，每个数字的高度为 0.60°，数字间间隔为 0.56°。掩蔽刺激的设计方式主要参考 Bhardwaj 等人的研究(2012)，分为以下两种：(a)一个白噪音矩形，每个噪音点的大小与目标刺激的笔画宽度相同，其中一半的噪音点的亮度与背景亮度相同，另一半与所设定的掩蔽刺激的亮度相同，两类噪音点随机分布，整个白噪音矩形大小在 5°视角内，可完全覆盖目标刺激；(b)与目标刺激位置和大小相同的随机正立汉字大写数字，字体为宋体，每个汉字数字均可完全覆盖对应的目标刺激。

所有刺激呈现在灰色背景上（RGB：(41,41,41)，亮度：16.0 cd/m^2），目标和掩蔽与背景的亮度对比分别为 50%和 100%。实验中采取高亮度对比的掩蔽刺激，以保证掩蔽刺激能够消除目标刺激的视觉后像带来的干扰，该亮度参数经预实验验证有效。5 名被试参加了预实验（男性 3 名，女性 2 名），平均年龄为 25.40±2.79 岁。所有被试均为在校大学生，母语均为汉语，均视力(或矫正视力)正常。被试在实验前不清楚实验目的。预实验采用与实验 1 完全相同的掩蔽刺激（随机白噪音和正立汉字大写数字）和目标刺激（阿拉伯数字），并且目标刺激与掩蔽刺激同时呈现。结果表明其绩效均处于猜测水平[1]，白噪音掩蔽条件下，被试绩效（1.18±0.11）与猜测水平（1.2）无显著差异（p=0.340）；正立汉字大写数字掩蔽下，被试绩效（1.30±0.11）与猜测水平（1.2）无显著差异（p=0.123），且在事后报告中所有被试均表示无法看清目标刺激、采用猜测方式作答。因此采用该亮度对比，能够保证掩蔽刺激具有足够的遮盖能力，避免正式实验中目标刺激的视觉后像带来的干扰。

2.1.3 实验设计与流程

实验主要采用经典的部分报告法任务，并与后掩蔽技术相结合，考察不同掩蔽刺激对被试记忆绩效的影响。实验为 2×3 被试内设计，每名被试的实验均包含两部分，分别采用两种不同的掩蔽类型（白噪音/汉字)。在每个部分的实验中均包含三个组块，分别对应三种不同的线索延迟时间：0ms、447ms 和无提示线索，其中无提示线索条件即为全部报告法条件，即被试需要报告全部 12 个目标阿拉伯数字。线索延迟时间指的是从

[1]可能的数字共 10 个，每个位置猜对的概率为 0.1，所有位置的猜测相互独立，将部分报告法结果折算为全部报告后，随机水平为 1.2 个。预实验中被试的报告绩效为 1.235 个。

掩蔽刺激消失后至线索呈现的时间，因而所有条件下对掩蔽刺激的视觉加工时间基本相同。采用组块设计的目的一方面是避免将任务要求相关的线索延迟时间作为组块内变量时，在线索呈现阶段带来复杂的决策过程；另一方面则是考虑到该任务难度较大，组块设计能够尽量提高被试绩效，避免绝对绩效过低而覆盖了掩蔽效应的差异。两种掩蔽类型和三种线索延迟时间的顺序通过拉丁方设计在被试间平衡。实验共包含 280 个试次，大约需要 60min，其中每个无提示线索条件组块包含 20 个试次，有线索条件组块则分别包含 60 个试次，以保证每个组块内任务所需报告的目标数相同。

部分报告任务的提示线索为 500ms 的提示音，其中 2.5kHz 高音，提示报告目标刺激的第一行；850Hz 中音，提示报告目标刺激的第二行；200Hz 低音，提示报告目标刺激的第三行。不同的提示线索在组块中随机平衡。

每个试次开始时，首先在屏幕中央呈现一个“+”（0.60°×0.60°）作为注视点， 1.4s 后注视点消失，随即呈现目标刺激 23.5ms（2 帧），目标刺激消失后空屏 105.9ms（9 帧），随后呈现掩蔽刺激 23.5ms（2 帧），掩蔽刺激消失后，根据不同的实验条件呈现提示线索，以指示被试需要报告的目标内容（图 1）。被试需要通过键盘按照从左到右、从上到下的顺序输入数字，屏幕上会显示被试所输入的数字，输入错误可以使用删除键重新输入，不确定时尽量猜测，完成输入后按回车键确认，直接进入下一个试次。

在正式实验开始前，被试需要进行一部分练习实验，该部分数据不纳入后续分析。练习实验包括两个组块：全部报告法组块（10 个试次）和 0ms 延迟的部分报告法组块（60 个试次），每个组块的一半试次为白噪音掩蔽，另一半试次为汉字掩蔽。

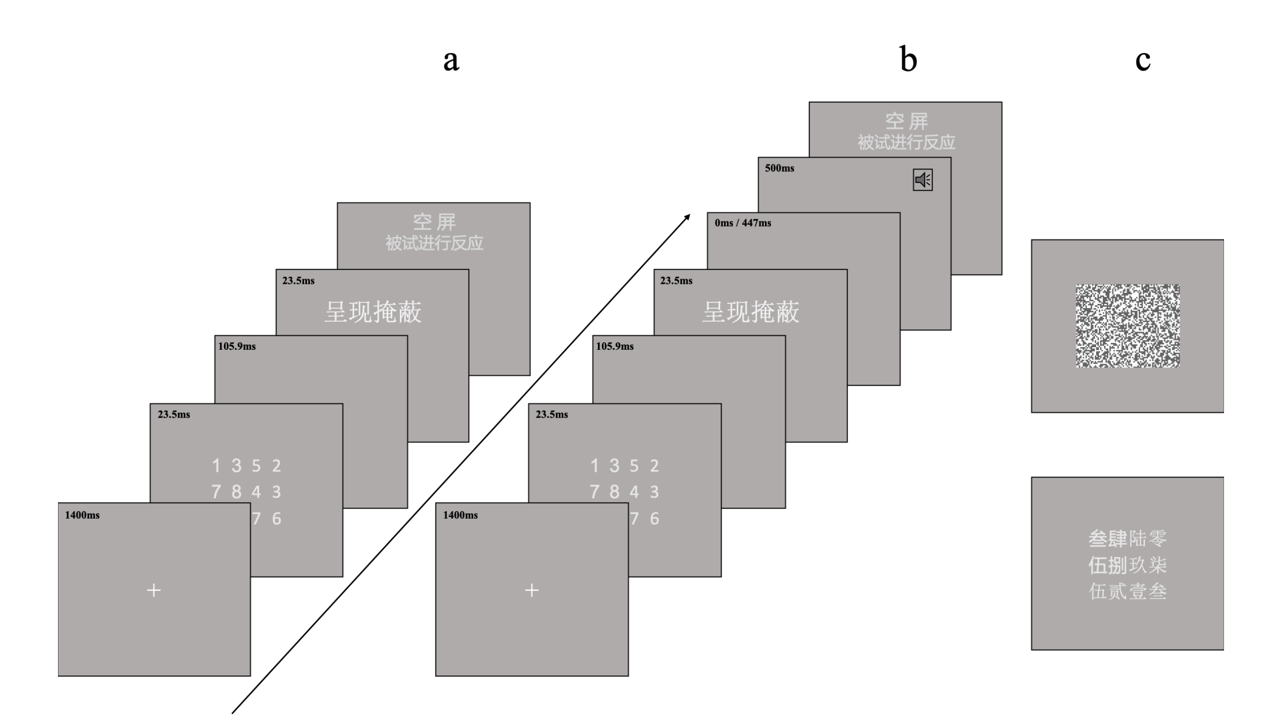


图 1 实验 1 流程图。掩蔽刺激消失后呈现空屏,被试根据不同的声音线索进行反应。(a) 在全部报告法组块中，不呈现声音线索;(b) 在部分报告法组块中，根据实验条件，在掩蔽刺激消失 0ms 或 447ms 后呈现声音线索提示被试需要报告哪一行 (2.5kHz 高音：报告第一行；850Hz 中音：报告第二行；200Hz 低音：报告第三行)。(c) 掩蔽刺激示例，白噪音掩蔽与正立汉字大写数字掩蔽。

2.2 结果与分析

实验中只有在对应位置上报告正确的数字才被记录为正确。在全部报告法中，被试的绩效是被试报告出正确数字的个数；在部分报告法中，被试只报告指定行中的数字，其绩效应为报告出的正确个数乘以行数（3 行）。

实验 1 中不同条件下被试的记忆绩效如图 2 所示。重复测量方差分析结果显示：掩蔽类型的主效应显著（$F(1,23) = 96.95$，$p < 0.001$，$\eta_p^2 = 0.81$），报告类型的主效应边缘显著（$F(2,46) = 3.10$，$p = 0.055$，$\eta_p^2 = 0.12$）；掩蔽类型与报告类型的交互效应不显著（$F(2,46) = 1.38$，$p = 0.262$）。

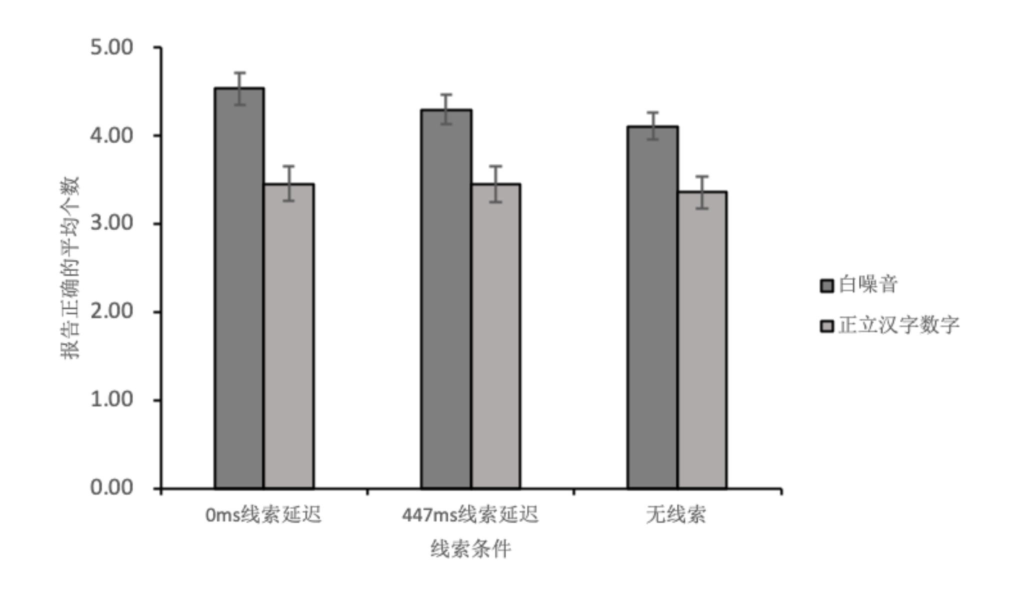


图 2 实验 1 的被试绩效。0ms 线索延迟与 447ms 线索延迟为部分报告法条件，无线索为全部报告法条件。

进一步检验（Bonferroni 校正）表明，在三种不同的报告类型中，白噪音掩蔽的被试绩效均显著优于正立汉字掩蔽（$p < 0.001$）。在白噪音掩蔽条件下，被试在 0ms 延迟条件下的绩效显著高于 447ms 延迟条件（$p = 0.023$）和无线索提示条件（$p = 0.009$），447ms 延迟条件与无线索提示条件的绩效间无显著差异（$p = 0.200$）；在正立汉字掩蔽条件下，被试在不同线索提示水平间的绩效均无显著差异（0ms 延迟 vs 447ms 延迟：$p = 0.968$； 0ms 延迟 vs 无线索提示：$p = 0.535$；447ms 延迟 vs 无线索提示：$p = 0.557$）。

在不同线索提示水平下稳定存在的掩蔽类型绩效差异表明，对于阿拉伯数字，正立汉字大写数字的掩蔽效果优于白噪音掩蔽，即汉字大写数字作为掩蔽刺激时与目标刺激产生了更大的冲突。考虑到汉字大写数字与阿拉伯数字具有完全不同的形态特征，仅存在语义信息方面的相似性，上述结果暗示，汉字大写数字与阿拉伯数字在视觉感觉记忆中存在语义信息的冲突，即感觉记忆中形成了带有语义信息的视觉表征。然而，尽管控制了汉字大写数字笔画粗细与白噪音掩蔽噪点大小相同，但汉字大写数字与白噪音的掩蔽差异也可能是由二者的外形差异所致。为了进一步确认语义相似性对掩蔽效应的影响，实验 2 将保持汉字大写数字的掩蔽刺激外形特点不变，利用倒置操作，针对性地破坏其语义信息，考察被试绩效的对应变化。

3 实验 2

3.1 方法

3.1.1 被试

共 24 名被试参与了实验 2（男性 8 人，女性 16 人），平均年龄为 22.04±2.44 岁；所有被试均为在校大学生，母语均为汉语，均视力（或矫正视力）正常；被试在实验前均不清楚实验目的。

3.1.2 刺激材料

除了掩蔽刺激外，本实验的刺激材料与实验 1 完全相同。掩蔽刺激分为以下两种：（a）与实验 1 相同的白噪音掩蔽；（b）倒置的汉字大写数字，其大小、位置等其他物理特征与实验 1 中正立汉字大写数字相同。（如图 3a）

3.1.3 实验设计与流程

采用倒置的汉字大写数字替换实验 1 中对应的正立汉字大写数字条件，其余实验设计和流程与实验 1 完全相同。

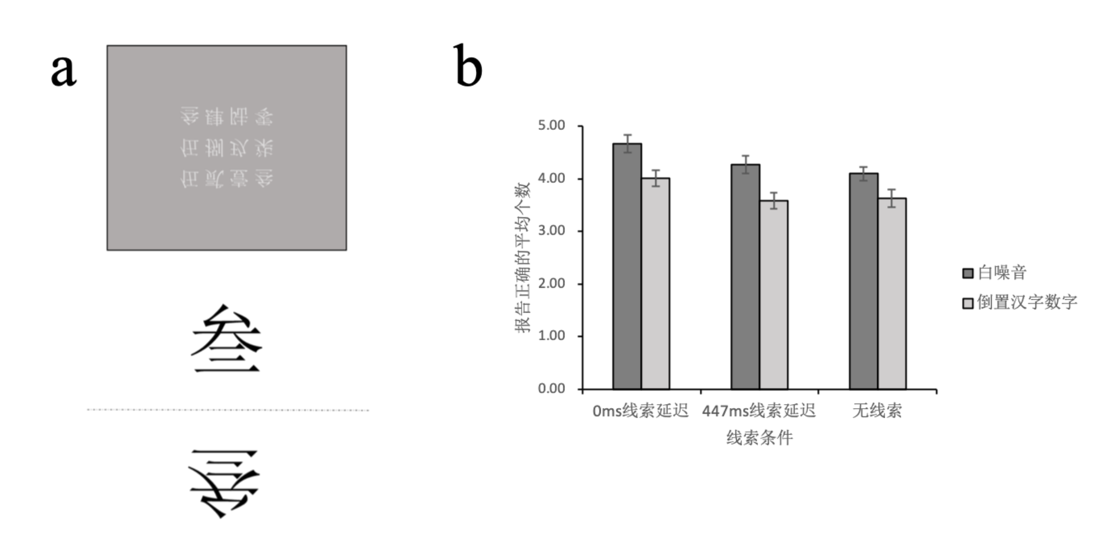


图 3 实验 2 刺激图片与被试绩效。（a）实验 2 中所用倒置汉字数字的示例；（b）实验 2 的被试绩效。

3.2 结果与分析

实验 2 中不同条件下被试的记忆绩效如图 3b 所示。重复测量方差分析结果显示：掩蔽类型的主效应显著（$F(1,23) = 63.55$，$p < 0.001$，$\eta_p^2 = 0.73$），报告类型的主效应显著（$F(2,46) = 15.09$，$p < 0.001$，$\eta_p^2 = 0.40$）；掩蔽类型与报告类型的交互效应不显著（$F(2,46) = 0.58$，$p = 0.564$）。

进一步检验（Bonferroni 校正）表明，在三种不同的报告类型中，白噪音掩蔽的被试绩效均显著优于倒置汉字掩蔽（0ms 延迟 & 447ms 延迟：$p < 0.001$；无线索提示：$p = 0.007$）。在白噪音掩蔽条件下，被试在 0ms 延迟条件下的绩效显著高于 447ms 延迟条

件（$p = 0.009$）和无线索提示条件（$p < 0.001$），447ms 延迟条件与无线索提示条件的绩效间无显著差异（$p = 0.135$）；在倒置汉字掩蔽条件下，被试在 0ms 延迟条件下的绩效显著高于 447ms 延迟条件（$p = 0.012$）和无线索提示条件（$p = 0.027$），447ms 延迟条件与无线索提示条件的绩效间无显著差异（$p = 0.803$）。

在不同线索提示水平下稳定存在的掩蔽类型绩效差异表明，倒置汉字大写数字的掩蔽效果同样优于白噪音掩蔽，这可能是由于两方面原因造成的：一方面，倒置汉字大写数字在形态结构上与白噪音具有较大差异，并且表现为离散的多个客体单元，更接近目标刺激的形态特征，即掩蔽刺激的外形特点对掩蔽效果存在影响；另一方面，倒置操作虽然抑制了汉字的语义信息提取，但仍有可能未完全破坏其语义信息，与白噪音完全不具有语义信息相比，倒置汉字大写数字可能残留的语义信息会对掩蔽效果存在影响。然而，实验 1 中正立汉字大写数字的掩蔽效应在不同的报告类型间无显著差异，即未表现出部分报告法效应。实验 2 中倒置汉字大写数字的掩蔽效应则在不同的报告类型间存在显著差异，即表现出部分报告法效应。这或许意味着，在外形特点对掩蔽效应存在影响的基础上，掩蔽刺激的语义信息同样会对掩蔽效应存在影响。

倒置操作对汉字语义信息的破坏是相对正立汉字而言的，直接对正立汉字大写数字与倒立汉字大写数字进行对比能够更具针对性地考察语义信息对掩蔽效应的影响。然而，考虑到对正立汉字大写数字的视觉体验，可能启动被试对倒置汉字大写数字语义信息的提取，本研究并未在实验内同时设置正立与倒置两类汉字大写数字掩蔽刺激。但鉴于实验 1 与实验 2 在实验环境、刺激材料、设计流程等方面高度相似，并存在相同的白噪音掩蔽条件，可采用跨实验的比较进行进一步检验。首先，笔者对实验 1 与实验 2 中白噪音掩蔽的三种线索条件分别进行独立样本 t 检验，结果显示，在三种线索条件下，实验 1 与实验 2 的白噪音被试绩效均无显著差异（0ms 延迟：$t(46) = -0.60$，$p = 0.579$；447ms 延迟：$t(46) = 0.14$，$p = 0.893$；无线索提示：$t(46) = 0.09$，$p = 0.93$）。为进一步检验无差假设，笔者进行了贝叶斯因子分析（Bayes Factor, BF）(吴凡，顾全，施壮华，高在峰，沈模卫，2018)，并使用适于 t 检验的柯西分布（Cauchy distribution）作为备择假设的先验(Jeffreys, 1998; Rouder, Speckman, Sun, Morey, & Iverson, 2009)。结果显示，$BF_{0ms\ 延迟} = 3.06$，$BF_{447ms\ 延迟} = 3.45$，$BF_{无线索提示} = 3.47$,三种线索条件下的贝叶斯因子均大于 3，即零假设成立的可能性是备择假设成立可能性的 3 倍以上，则有足够强的证据提示实验 1 与实验 2 的白噪音的掩蔽无差异。其次，笔者对实验 1 与实验 2 中汉字掩蔽的三种线索提示条件分别进行独立样本 t 检验，结果显示，在 0ms 延迟条件下，倒置汉字掩蔽下的被

试绩效显著高于正立汉字（$t(46) = -2.21$，$p = 0.032$，*Cohen'd* = 1.37）；在 447ms 延迟和无线索提示条件下，二者无显著差异（447ms 延迟：$t(46) = -0.53$，$p = 0.600$；无线索提示：$t(46) = -1.13$，$p = 0.263$）。

上述分析结果显示，两组被试在白噪音掩蔽条件下的绩效无显著差异，表明两个实验具有同质性；同时，正立汉字大写数字的掩蔽效果较倒置汉字大写数字更强，表明掩蔽刺激的语义信息对掩蔽效果存在影响，即感觉记忆的视觉表征中包含语义信息。然而，倒置汉字数字几乎不具有语义信息，正立汉字数字与倒置汉字数字的掩蔽效应差异可能源于掩蔽刺激是否具有语义信息，而非掩蔽刺激与目标刺激的语义相似性。这种可能性会削弱关于感觉记忆的视觉表征中包含语义信息的结论。为进一步确认语义相似性对掩蔽效应的影响，实验 3 将选用与汉字大写数字字频和笔画数相似的汉字常用字作为对照掩蔽刺激，直接比较汉字大写数字与汉字常用字的掩蔽效应。

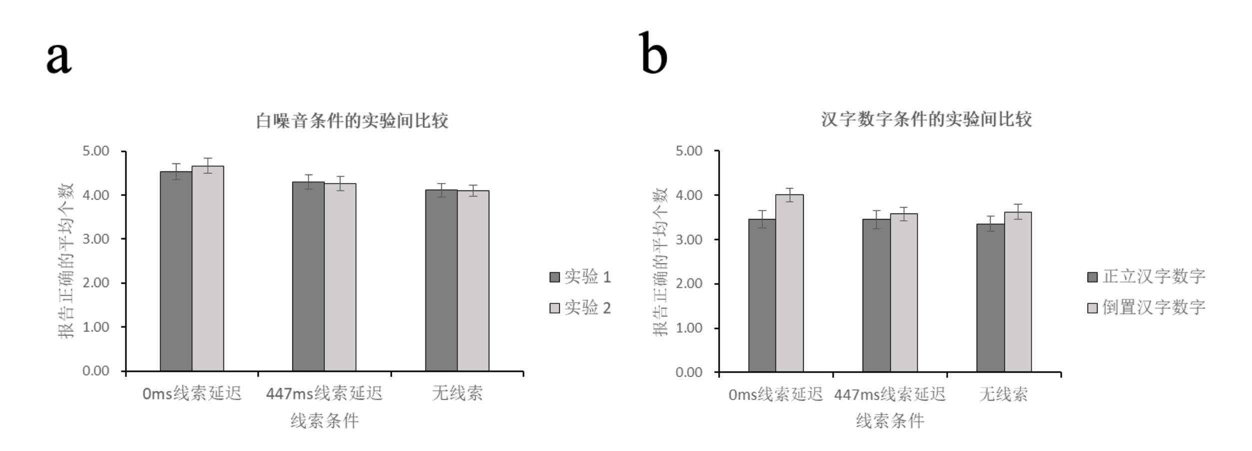


图 4 实验 1 与实验 2 的被试绩效比较。(a) 实验 1 与实验 2 的白噪音掩蔽条件比较；(b) 实验 1 与实验 2 的汉字数字掩蔽条件比较。

4 实验 3

4.1 方法

4.1.1 被试

共 24 名被试参与了实验 3（男性 10 人，女性 14 人），平均年龄为 19.04±1.52。被试均为在校大学生，母语均为汉语，均视力（或矫正视力）正常；被试在实验前均不清楚实验目的。

4.1.2 刺激材料

除了掩蔽刺激外，本实验的刺激材料与实验 1 完全相同。掩蔽刺激分为以下两种：（a）正立的汉字大写数字；（b）正立的汉字常用字。汉字呈现的大小、位置等其他物理特征与实验 1、2 中的汉字大写数字相同。根据字频(现代汉语频率词典，1986)与笔画数选用汉字常用字如下：柬、筏、蓉、稚、休、胡、岔、淑、沧、香，以保证所选常用字与汉字大写数字在物理特性和熟悉度方面尽量相似。汉字常用字的字频为 0.00409%±0.00624%，汉字大写数字的字频为 0.00411%±0.00631%；汉字常用字的笔画数为 9.70±2.45，汉字大写数字的笔画数为 9.40±2.55。

4.1.3 实验设计与流程

采用汉字常用字替换实验 1 中对应的白噪音条件，其余实验设计和流程与实验 1 完全相同。

4.2 结果与分析

实验 3 中不同条件下被试的记忆绩效如图 5 所示。重复测量方差分析结果显示：掩蔽类型的主效应显著（$F(1,23) = 4.53$，$p = 0.044$，$\eta_p^2 = 0.17$），报告类型的主效应显著（$F(2,46) = 5.47$，$p = 0.007$，$\eta_p^2 = 0.19$）；掩蔽类型与报告类型的交互效应不显著（$F(2,46) = 1.99$，$p = 0.148$）。

进一步检验（Bonferroni 校正）表明，在三种不同的报告类型中，在 0ms 延迟条件下，汉字常用字掩蔽的被试绩效显著优于汉字大写数字掩蔽（$p = 0.024$）；在 447ms 延迟与无线索提示条件下，被试绩效在两种掩蔽条件下无显著差异（447ms 延迟：$p = 0.110$；无线索提示：$p = 0.661$）。在汉字常用字掩蔽条件下，被试在 0ms 延迟（$p = 0.003$）和 447ms（$p = 0.017$）延迟条件下的绩效均显著高于和无线索提示条件，0ms 延迟与 447ms 延迟条件的绩效间无显著差异（$p = 0.576$）；在汉字大写数字下，被试在不同线索提示水平间的绩效均无显著差异（0ms 延迟 vs 447ms 延迟：$p = 0.892$； 0ms 延迟 vs 无线索提示：$p = 0.433$；447ms 延迟 vs 无线索提示：$p = 0.524$）。

上述结果显示，具有同样强度语义信息的情况下，与目标刺激具有更高语义相似性的汉字大写数字，相较汉字常用字具有更强的掩蔽效应，这表明掩蔽刺激与目标刺激的语义相似性影响掩蔽效应，即感觉记忆的视觉表征中包含语义信息。

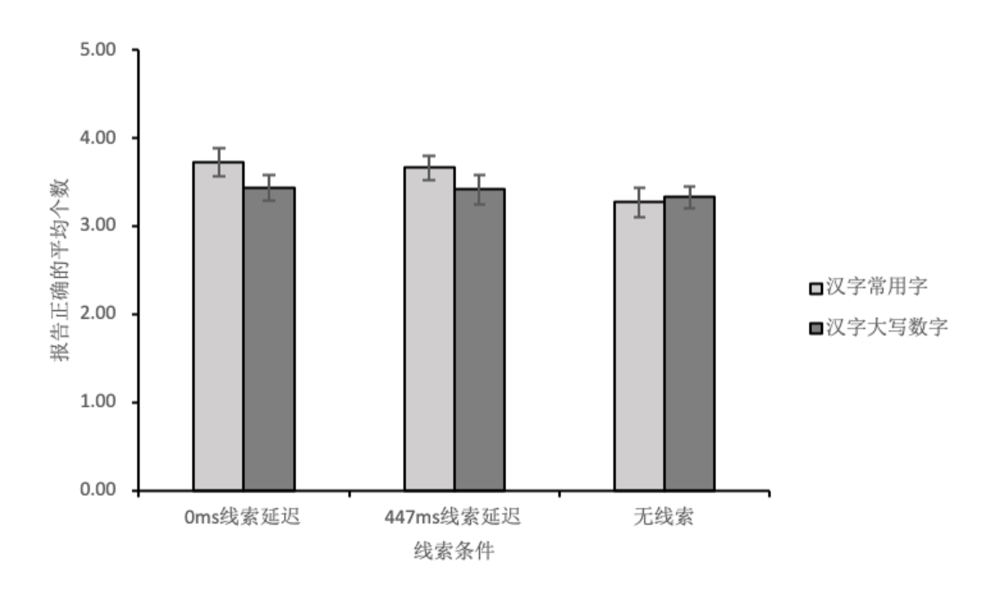


图 5 实验 3 的被试绩效

5 总讨论

本研究结合部分报告法与掩蔽技术，通过三个实验考察了抽象信息在感觉记忆中的视觉表征。实验 1 发现，与不带有语义信息的白噪音掩蔽相比，含语义信息的正立汉字大写数字具有更强的掩蔽效应。实验 2 通过倒置汉字破坏其语义信息，尽管倒置汉字的掩蔽效应仍较白噪音掩蔽较强，但进一步的实验间比较表明正立汉字与倒置汉字相比具有更强的掩蔽效应。实验 3 采用与正立汉字大写数字具有相似物理特性和熟悉度的汉字常用字作为掩蔽刺激，结果表明与目标刺激具有更高语义相似性的正立汉字大写数字，其掩蔽效应较汉字常用字更强。掩蔽效应的增强（即记忆绩效的降低）源于感觉记忆中对目标刺激与掩蔽刺激的表征冲突(Bhardwaj et al., 2012; Pinto et al., 2013; Smithson & Mollon, 2006; Zerr et al., 2017)，可以推论，感觉记忆的视觉表征中包含语义信息。

采用部分报告法对感觉记忆进行探究时往往设计多档线索延迟时间，并与全部报告法相比较，由线索延迟时间带来的差异被称为部分报告法效应。本研究主要关注不同类型掩蔽刺激对感觉记忆的影响，因而注重分析不同掩蔽类型条件下的绝对绩效差异，然而对部分报告法效应进行一定的分析仍有助于完善本研究的论证。部分报告法任务中通常采用键盘、纸笔或口头的方式进行作答，无论采用何种作答方式，反应过程总是需要工作记忆的参与，因而部分报告法效应的存在保证了实验范式所探究的确实是感觉记忆的加工情况。本研究的三个实验中，线索延迟时间的主效应均达到显著水平，表明存在部分报告法效应，即研究揭示的是视觉感觉记忆的加工情况。此外，部分报告法效应的

强弱一定程度上也能够反应掩蔽效应的强弱，强的掩蔽效应会造成线索延迟较短时也无法达到较高的绩效，表现出部分报告法效应减弱。本研究中白噪音掩蔽均具有较强的部分报告法效应，倒立汉字大写数字与常用字作为掩蔽时，尽管绝对绩效降低，但仍表现出了部分报告法效应，而正立汉字大写数字作为掩蔽时，完全破坏了部分报告法效应，这同样表明正立汉字大写数字具有更强的掩蔽效应，即感觉记忆的视觉表征中包含语义信息。

表征是视觉领域研究的核心问题之一。早期观点认为视觉感觉记忆以快照形式构建视觉表征，快照形式表征的主要特点是相继呈现的刺激的外形特征，会影响对刺激的内在表征。本研究中所有汉字掩蔽的外形结构和纹理特征均与白噪音存在较大差异，且均表现出较白噪音刺激更强的掩蔽效应，一定程度上说明在视觉加工初期，快照式的表征的确存在。

与早期的快照式表征观点不同，近年来有研究者提出了特征形式的表征观点(Bhardwaj et al., 2012)。该观点认为，视觉感觉记忆能够提取刺激不同维度的特征，并以相对抽象的特征形式构建视觉表征。在实验中表现为目标刺激与掩蔽刺激在物理形式上不同但存在特征相似性时，仍然能够产生表征的冲突，导致掩蔽效应提升。本研究的结果同样支持特征形式表征的存在，一方面实验中目标刺激是阿拉伯数字，和汉字在外形特征方面不尽相同，然而汉字掩蔽刺激仍能对阿拉伯数字产生较强的掩蔽效应；另一方面实验中的正立汉字大写数字、倒置汉字大写数字和汉字常用字均具有高度相似的外形特征，然而正立汉字大写数字与目标刺激的语义相似性仍表现出显著更强的掩蔽效应。这两方面证据表明，相继呈现刺激之间的掩蔽既依赖于刺激的外形特征，还依赖于刺激的抽象特征（如：语义信息），即感觉记忆中的语义信息与刺激具体形象分离，能够以抽象的特征形式构建对刺激的表征。

考察感觉记忆特征形式表征的以往研究往往采用相对简单、底层的特征（如：朝向、外形结构等）(Bhardwaj et al., 2012; Pinto et al., 2013)，这类特征尽管在感觉记忆的视觉表征中可以以抽象形式存在，但对刺激本身特征信息的提取，极大地依赖于刺激表现出的物理信息。与之不同，本研究所采用的语义信息本身属抽象信息，对其的提取依赖于个体的知识经验(Federmeier & Kutas, 1999)，反映了认知加工中先验信息的作用。本研究对感觉记忆视觉表征中是否包含语义信息的针对性考察，有助于完善感觉记忆表征的相关理论。

感觉记忆的表征中包含刺激的抽象信息这一特点具有重要的功能意义。人类智能的

核心体现之一是其“小样本、大任务”的学习和工作特性(唐宁，安玮，周吉帆，高涛，沈模卫，2018; 周吉帆 等，2016)，为实现该特性，人类的认知系统需要将由学习得到的知识经验迁移到不同场景，并在后续的认知加工中体现出来。包含语义信息的视觉感觉记忆表征为不同认知模块的交互提供了可能，表明人类的先验知识在视觉加工的最初阶段就已产生作用，并能够持续参与到整个视觉加工过程中。此外，感觉记忆的核心功能是在选择性加工发生之前对短暂呈现的大量刺激进行存储(Öğmen & Herzog, 2016; Smithson & Mollon, 2006)，由于进一步的选择性加工并非仅基于刺激的外在物理特性，往往还与刺激的抽象信息、个体的任务等因素相关(Wahn & König, 2017)，仅依赖于快照形式的表征可能对视觉加工带来阻碍(Haber, 1983)。因此感觉记忆的表征中包含刺激的抽象信息，为进一步的选择加工提供可能。

需要指出的是，本研究并不否认基于快照形式的表征观点。实验中倒置汉字的掩蔽效应仍强于白噪音，尽管倒置操作可能未完全破坏汉字的语义信息，然而更为重要的因素在于倒置汉字的形态结构特征与白噪音具有较大差异，其纹理分布更为复杂、与整块的白噪音相比更可能被识别为多个分离的客体、形态上与目标刺激更为接近，这表明感觉记忆中包含了以快照形式存在的对刺激对象物理特性的表征。同时正立汉字大写数字与倒置汉字大写数字掩蔽效应的实验间比较，以及正立汉字大写数字与汉字常用字掩蔽效应的实验内比较，均表明感觉记忆中还包含了以特征形式存在的对刺激对象抽象信息的表征。基于此，本研究认为快照形式与特征形式可能共同构成了感觉记忆的视觉表征。快照式的表征与视知觉表征更为接近，依赖于空间位置信息(Dick, 1974)。特征形式的表征则与视觉工作记忆的表征更为接近，其中物理位置对对象的限制作用下降(宋旭晨，2017)。视觉感觉记忆的表征在时间维度上是否表现出由快照形式向特征形式的转换，仍有待进一步研究探讨。

6 小结

本研究通过测量具有不同语义信息的掩蔽刺激对相同目标刺激的掩蔽效应差异，考察了感觉记忆的视觉表征特点，获得以下结论：感觉记忆的视觉表征中包含了刺激对象的语义信息，包含语义信息的视觉表征以特征形式存在，并可能与快照式的表征共存。

参考文献

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *The Bell System Technichal Jounarl*, *40*(1), 309–328.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1–10.

Bhardwaj, R., Mollon, J. D., & Smithson, H. E. (2012). Compatibel and incompatible representation in visual sensory storage. *Journal of Vision*, *12*(5), 1–10.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Coltheart, M., Laming, D. R. J., Routh, D. A., & Broadbent, D. E. (1983). Iconic memory [and discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *302*(1110), 283–294.

Dick, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. *Perception & Psychophysics*, *16*(3), 575–596.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469–195.

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 845–866.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *The Behavioral and Brain Sciences*, *6*, 1–54.

Jeffreys, H. (1998). *Theory of probability* (3rd ed.). *Oxford classic texts in the physical sciences*. Oxford Oxfordshire, New York: Clarendon Press; Oxford University Press.

Kao, C.-H., Chen, D.-Y., & Chen, C.-C. (2010). The inversion effect in visual word form processing. *Cortex*, *46*(2), 217–230.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjuctions. *Nature*, *390*(6657), 279–281.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, Oxford: Freeman.

Modern Chinese frequency dictionary (in Chinese). (1986). Beijing: Beijing Language Institute Press.

[*现代汉语频率词典*. (1986). 北京：北京语言学院出版社]

Rescorla, M. (2015). Bayesian perceptual psychology. M. Matthen(Ed.), *The Oxford handbook of philosophy of perception* (First edition) (pp. 694–716). Oxford, New York NY: Oxford University Press.

Neisser, U. (1967). *Cognitive psychology*. *The Century psychology series*. New York: Appleton-Century-Crofts.

Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience*, *9*(9), 1186–1192.

Öğmen, H., & Herzog, M. H. (2016). A new conceptualization of human visual sensory-memory. *Frontiers in Psychology*, *7*, 1–15.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, *30*(1), 473–485.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 633–643.

Pinto, Y., Sligte, I. G., Shapiro, K. L., & Lamme, V. A. F. (2013). Fragile visual short-term memory is an object-based and location-specific store. *Psychonomic Bulletin & Review*, *20*(4), 732–739.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an rsvp task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860.

Rensink, R. A. (2002). Change detection. *Annual Reviews*, *53*, 245–277.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237.

Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. *Psychologische Forschung*, *36*(1), 71–93.

Schlottmann, A., & Ray, E. (2010). Goal attribution to schematic animals: Do 6-month-olds perceive biological motion as animate? *Developmental Science*, *13*(1), 1–10.

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. M.D. Rutherford & Valerie A. Kuhlmeier (Eds). *Social perception: Detection and interpretation of animacy, agency, and intention* (pp.197–230). Cambridge, Massachusetts: MIT Press.

Schultz, D. W., & Eriksen, C. W. (1977). Do noise masks terminate target processing? *Memory & Cognition*, *5*(1), 90–96.

Shen, M. W., Gao, T., & Ding, H. J. (2004). A study of Chinese and Arabic numerals' subliminal priming effects. *Journal of Psychological Science*, *27*(1), 13–17.

[沈模卫，高涛，丁海杰. (2004). 汉字数字与阿拉伯数字的阈下启动研究. *心理科学*, 27(1), 13–17.]

Shen, M., Pan, S., & Chen, X. (1998). A study on the cognitive process of identification of two–radical Chinese characters. *Chinese Journal of Applied Psychology*, *4*(1), 27–32.

[沈模卫，潘善会，陈新. (1998). 合体字字形识别过程探索. *应用心理学*, *4*(1), 27–32.]

Smithson, H., & Mollon, J. (2006). Do masks terminate the icon? *Quarterly Journal of Experimental Psychology*, *59*(1), 150–160.

Song, X. C. (2017). Grouping functionally-related objects in working memory (Unpublished master's thesis). Zhejiang University, Hangzhou.

[宋旭晨. (2017). *功能相关的客体在工作记忆中的组织*(硕士学位论文). 浙江大学，杭州.]

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors*, *5*(1), 19–31.

Tang, N., An, W., Zhou, J. F., Gao, T., & Shen, M. W. (2018). From "data" to "representation": How human cognition can inspire artificial intelligence. *Chinese Journal of Applied Psychology*, *24*(1), 3–14.

[唐宁，安玮，周吉帆，高涛，沈模卫. (2018). 从数据到表征：人类认知对人工智能的启发. *应用心理学*, *24*(1), 3–14.]

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, *80*(1), 1–52.

Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task-dependent? *Advances in Cognitive Psychology*, *13*(1), 83–96.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 483–502.

Wu, F., Gu, Q., Shi, Z. H., Gao, Z. F., & Shen, M. W. (2018). Striding over the "classical statistical inference trap": Application of Bayes Factors in psychological studies. *Chinese Journal of Applied Psychology*, *24*(3), 195–202.

[吴凡，顾全，施壮华，高在峰，沈模卫. (2018). 跳出传统假设检验方法的陷阱——贝叶斯因子在心理学研究领域的应用. *应用心理学*. *24*(3), 195–202.]

Yin, J., & Csibra, G. (2015). Concept-based word learning in human infants. *Psychological Science*, *26*(8), 1316–1324.

Zerr, P., Gayet, S., Mulder, K., Pinto, Y., Sligte, I., & van der Stigchel, S. (2017). Remapping high-capacity, pre-attentive, fragile sensory

memory. *Scientific Reports*, *7*(1), 1–10.

Zhou, J. F., Xu, H. K., Tang, N., Shi, B. H., Zhao, Y., Gao, T., & Shen, M. W. (2016). The “strong-cognition" sychology research: Implication of AlphaGo. *Chinese Journal of Applied Psychology*, *22*(1), 3–11.

[周吉帆，徐昊骙，唐宁，史博皓，赵阳，高涛，沈模卫. "强认知"的心理学研究：来自 AlphaGo 的启示. *应用心理学*. *22*(1), 3–11.]

Visual Representation of Items with Semantic Information in Sensory Memory

HU Jingjing[1]; XU Haokui[1]; CAO Liren[1]

([1]Department of Psychology and Behavioral Science, Zhejiang University, Hangzhou 310028, China)

Abstract

Visual representation is one of the most important topics in sensory memory. Traditionally, visual representation with respect to sensory memory was regarded as an 'icon', the processing of which would cease after a subsequent stimulus (backward masking). Recently, an increasing number of researches have suggested a more flexible representation of visual sensory memory in which the features of the visual stimulus could be expressed in an abstract form. As previous researchers paid considerable attention to only the schematic features of a visual stimulus, the abstract information contained in a stimulus has rarely been explored. In this study, we have focused on the semantic information of the visual representation of the sensory memory. In a part-report task, the Arabic digits were employed as the target, and the Chinese digits as the semantic-similar mask, which were semantically related to Arabic digits; the noise patterns were employed as the non-semantic mask. If under the Chinese mask condition, the participants' performance of reporting target digits decreased, it would reveal a conflict between the representations of the target and the mask, thereby suggesting that the visual representation of the sensory memory contains semantic information of the stimulus.

In Experiment 1, the target stimuli were arrays of 3×4 Arabic digits generated by selecting at random from a set containing digits zero to nine. A mask stimuli could either be (a) a single binary visual noise with an element size equal to the stroke width of the target digits, or (b) a 3×4 array of random standard Chinese-digits in capital form whose spatial arrangement is identical to that of the target and its size is big enough to overlap the corresponding target digits. Both the targets and masks were presented within 23.5 ms (2 frames in 85 Hz) and the inter-stimulus interval (ISI) between the offset of the target and the onset of the mask was 105.9 ms (9 frames in 85 Hz). There were three conditions for reporting: part-report with 0 ms cue-delay, part-report with 447 ms cue-delay, and whole-report with no cue-delay. A correct response was recorded only when the participants reported the correct

digits at the locations they occurred in the target array. The results showed a significant difference between the noise and standard Chinese-digit mask under all of the three reporting conditions; there was a significant part-report advantage for the noise mask, but not for the Chinese characters. This indicated that semantic information can influence information processing in the sensory memory. Considering that noise and Chinese characters are different at a schematic level, we inverted the Chinese digits in Experiment 2 to destroy their semantic information but reserve their schematic feature.

Therefore, Experiment 2 was identical to Experiment 1, except the Chinese digits were inverted. The results showed that the difference between the noise and inverted Chinese-digit masks still existed, while both showed significant part-report advantages. We conducted an analysis across-experiment afterwards. There was a significant difference between the standard and inverted Chinese-digit masks, while there was no significant difference between the noise masks across the two experiments. Results above revealed that masks with different semantic information but the same schematic information can influence the target processing differently. Considering that the standard and inverted Chinese digits are different in semantic strength, we employed commonly used Chinese characters in Experiment 3 to ensure that the two masks under comparison have different meanings but a similar semantic strength.

In Experiment 3, we employed Chinese digits and commonly used Chinese characters as masks. These two masks had similar word frequencies and number of strokes. However, the Chinese-digit mask still showed a significant stronger masking effect, which revealed the effect of semantic similarity. This implied that the visual representation in sensory memory could contain semantic information.

In conclusion, our findings demonstrate that the feature represented in sensory memory is not only schematic but also semantic. The functions of such a feature-based representation and its interaction with the icon have also been discussed.

Keywords: sensory memory; visual representation; semantic information; masking